JOURNAL OF MULTIVARIATE ANALYSIS 31, 187–200 (1989)

Maximum Likelihood Estimators in Multivariate Linear Normal Models*

DIETRICH VON ROSEN

University of Stockholm, Stockholm, Sweden

A unified approach of treating multivariate linear normal models is presented. The results of the paper are based on a useful extension of the growth curve model. In particular, the finding of maximum likelihood estimators when linear restrictions exist on the parameters describing the mean in the growth curve model is considered. The problem with missing observations is also discussed and the EM algorithm is applied. Furthermore, a multivariate covariance model is generalized. © 1989 Academic Press, Inc.

1. INTRODUCTION

The main theme of this paper is to derive maximum likelihood estimators for parameters in multivariate linear normal models in three different situations: (i) no restrictions exist on the mean, (ii) restrictions exist, and (iii) missing observations exist (unrestricted and restricted case). To be more specific, the models considered are the following two;

DEFINITION 1. Multivariate linear normal model with mean ABC, referred to as MLNM(ABC). Let X: $p \times n$, A: $p \times q$ $q \leqslant p$, B: $q \times k$, C: $k \times n$ $\rho(C) + p \leqslant n$ and Σ: $p \times p$ p.d. The columns of X are independently p-variate normally distributed with an unknown dispersion matrix Σ and $E[X] = A\mathrm{B}C$, where A and C are known design matrices and B is an unknown parameter matrix.

DEFINITION 2. Multivariate linear normal model with mean $\sum_{i=1}^{m} A_i \mathrm{B}_i C_i$, referred to as MLNM($\sum_{i=1}^{m} A_i \mathrm{B}_i C_i$). Let X: $p \times n$, A_i: $p \times q_i$ $q_i \leqslant p$, B_i: $q_i \times k_i$, C_i: $k_i \times n$, Σ: $p \times p$ p.d., where $\rho(C_1) + p \leqslant n$ and $\mathbf{R}(C_m') \subseteq \mathbf{R}(C_{m-1}') \subseteq \cdots \subseteq \mathbf{R}(C_1')$. The columns of X are independently p-variate normally distributed with an unknown dispersion matrix Σ and

Received April 13, 1987; revised May 23, 1988.

AMS 1980 subject classifications: Primary 62H12; Secondary 62A10.

Keywords and phrases: growth curve model, linear restrictions, missing data, covariance model.

* Dedicated to the students in Peking.

0047-259X/89 $3.00

$E[X]=\sum_{i=1}^{m} A_i\mathrm{B}_iC_i$ where the A's and C's are known and the B's unknown.

Note that in the definitions as well as in the subsequent $\rho(\cdot)$ and $\mathbf{R}(\cdot)$ denote the rank and range space, respectively, and p.d. stands for positive definite. It is obvious that the MLNM(ABC) is a special case of the MLNM($\sum_{i=1}^{m} A_i\mathrm{B}_iC_i$). On the other, the MLNM(ABC) with certain linear restrictions on B is equivalent to the MLNM($\sum_{i=1}^{m} A_i\mathrm{B}_iC_i$) and this is utilized in Section 3.

The MLNM(ABC) was introduced by Potthoff and Roy [17], though some other authors had earlier treated special cases. Usually the model is called the growth curve model, GMANOVA model, generalized linear model, or just the Potthoff and Roy model. For various aspects of estimating B and Σ, the reader is referred to Potthoff and Roy [17], Khatri [12], Rao [18, 19], Gleser and Olkin [8], Grizzle and Allen [9], and Srivastava and Khatri [26]. For some good reviews of the model we suggest the work by Woolson and Leeper [28] and Seber [24]. An up to date review of the model is also given by von Rosen [21]. Two further works which have to be mentioned are: the book written by Kariya [11], where many useful results for testing hypotheses are collected, and the paper by Anderson and Olkin [1], where the problem of estimating parameters for the MLNM(ABC) in the unrestricted and restricted cases are thoroughly discussed.

The MLNM($\sum_{i=1}^{m} A_i\mathrm{B}_iC_i$) was introduced by von Rosen [20] and a canonical form of the model has been considered by Banken [4] (see also Chapter 4 in Kariya's book). A similar extension of the MLNM(ABC) was also indicated by Elswick [7].

The main purpose with this paper is to utilize some ideas presented by von Rosen [20] so that estimators for the parameters in the MLNM($\sum_{i=1}^{m} A_i\mathrm{B}_iC_i$) are obtainable (Section 2) and then present estimators for B in the MLNM(ABC) when linear restrictions exist such as $D\mathrm{B}E=0$ or $D_1\mathrm{B}E_1=0$ $D_2\mathrm{B}E_2=0$, or more general (Section 3). Note that hitherto such estimators have not been presented although Anderson and Olkin [1] indicate how to derive them in a canonical version of the MLNM(ABC; $D\mathrm{B}E=0$) and Banken [4] also in the canonical form (more general case) derives estimators recursively which, however, are rather hard to interpret. Moreover, when dealing with missing observations (Section 4) it is intended to extend the work by Liski [15] so that estimators in the MLNM(ABC) with restrictions on B can be derived. Thus, an alternative approach to Kleinbaum's [13] approach for estimating parameters and testing hypothesis when missing observations exist is presented. In Section 5 we make some notes on an extension of the MLNM($\sum_{i=1}^{m} A_i\mathrm{B}_iC_i$) which allows covariables in the model. However, all results follow immediately from the previous sections and therefore the model is not treated as a separate model. On the other hand, the results extend some of those given by Chinchilli and Elswick [5], Elswick [7], and von Rosen [20, 23].

2. Deriving ML-Estimators for the MLNM($\sum_{i=1}^{m} A_i \mathbf{B}_i C_i$)

Let, for arbitrary matrices G and F, G^- denote an arbitrary g-inverse in the sense of $GG^-G=G$ and F^0 any matrix spanning $\mathbf{R}(F)^\perp$, the orthogonal complement to $\mathbf{R}(F)$. In this section we are going to prove the following theorem.

Theorem 2.1. *Let*

$$P_r = T_{r-1}T_{r-2}T_{r-3} \times \cdots \times T_0, \qquad T_0 = I, \qquad r = 1, 2, ..., m+1$$

$$T_i = I - P_iA_i(A_i'P_i'S_i^{-1}P_iA_i)^- A_i'P_i'S_i^{-1}, \qquad i = 1, 2, ..., m$$

$$S_i = \sum_{j=1}^{i} K_j, \qquad i = 1, 2, ..., m$$

$$K_j = P_jXC_{j-1}'(C_{j-1}C_{j-1}')^- C_{j-1}(I - C_j'(C_jC_j')^- C_j) \\ \times C_{j-1}'(C_{j-1}C_{j-1}')^- C_{j-1}X'P_j', \qquad C_0 = I.$$

Assuming S_1 to be p.d., representations of the maximum likelihood estimators for the MLNM($\sum_{i=1}^{m} A_i \mathbf{B}_i C_i$) *are given by*

$$\hat{\mathbf{B}} = (A_r'P_r'S_r^{-1}P_rA_r)^- A_r'P_r'S_r^{-1}\left(X - \sum_{i=r+1}^{m} A_i\hat{\mathbf{B}}_iC_i\right)C_r'(C_rC_r')^- \\ + (A_r'P_r')^0 Z_{r1} + A_r'P_r'Z_{r2}C_r^{0'} \qquad r = 1, 2, ..., m$$

$$n\hat{\Sigma} = \left(X - \sum_{i=r+1}^{m} A_i\hat{\mathbf{B}}_iC_i\right)\left(X - \sum_{i=r+1}^{m} A_i\hat{\mathbf{B}}_iC_i\right)' \\ = S_m + P_{m+1}XC_m'(C_mC_m')^- C_mX'P_{m+1}',$$

where the Z's are arbitrary matrices.

Proof. The proof will utilize the following two lemmas of which the first is a well-known matrix relation and the second specific to the theorem.

Lemma 2.1. *Let S be a non-singular and V an arbitrary matrix, of proper dimensions. Then*

(i) $(S+VV')^{-1}=S^{-1}-S^{-1}V(V'S^{-1}V+I)^{-1}V'S^{-1}$

(ii) $(S+VV')^{-1}V=S^{-1}VH$

where H p.d. is obtainable from (i).

LEMMA 2.2. *Let P_r and S_r be defined as in Theorem* 2.1;

$$P_r'S_r^{-1}P_r=S_r^{-1}P_r, \qquad P_rP_r=P_r, \qquad r=1,2,...,m.$$

Proof. We will use an induction argument. Note that the lemma is true for $r=1,2$. Suppose that $P_{r-1}'S_{r-1}^{-1}P_{r-1}=S_{r-1}^{-1}P_{r-1}$ and $P_{r-1}P_{r-1}=P_{r-1}$. By applying Lemma 2.1(i) to $S_r^{-1}=(S_{r-1}+K_r)^{-1}$ it follows that it is enough to show that $P_r'S_{r-1}^{-1}P_r=S_{r-1}^{-1}P_r=P_r'S_{r-1}^{-1}$. However,

$$P_r'S_{r-1}^{-1}P_r=P_{r-1}'S_{r-1}^{-1}P_r=S_{r-1}^{-1}P_{r-1}P_r=S_{r-1}^{-1}P_r$$

and

$$P_rP_r=T_{r-1}P_{r-1}P_r=T_{r-1}P_r=P_r. \quad \blacksquare$$

Now we are going to prove the theorem by aid of the above lemmas. It is easy to see that the $m+1$ likelihood equations are equivalent to

$$A_r'\Sigma^{-1}\left(X-\sum_{i=1}^{m}A_i\mathrm{B}_iC_i\right)C_r'=0, \qquad r=1,2,...,m \tag{2.1}$$

$$n\Sigma=\left(X-\sum_{i=1}^{m}A_i\mathrm{B}_iC_i\right)\left(X-\sum_{i=1}^{m}A_i\mathrm{B}_iC_i\right)'. \tag{2.2}$$

Starting with $r=1$, we obtain the equations

$$A_1'\Sigma^{-1}(XC_1'(C_1C_1')^{-}C_1-\sum_{i=1}^{m}A_i\mathrm{B}_iC_i)C_1'=0 \tag{2.3}$$

$$n\Sigma=S_1+\left(XC_1'(C_1C_1')^{-}C_1-\sum_{i=1}^{m}A_i\mathrm{B}_iC_i\right)\times\left(XC_1'(C_1C_1')^{-}C_1-\sum_{i=1}^{m}A_i\mathrm{B}_iC_i\right)'. \tag{2.4}$$

Using Lemma 2.1(ii) shows that (2.3) can be written as

$$A_1'S_1^{-1}\left(XC_1'(C_1C_1')^{-}C_1-A_1\mathrm{B}_1C_1-\sum_{i=2}^{m}A_i\mathrm{B}_iC_i\right)HC_1'=0. \tag{2.5}$$

However, since $\mathbf{R}(C'_m) \subseteq \mathbf{R}(C'_{m-1}) \subseteq \cdots \subseteq \mathbf{R}(C'_1)$ and H is p.d. we have $\mathbf{R}(C_i H C'_1) = \mathbf{R}(C_i)$ meaning that (2.5) is independent of H and therefore a consistent linear equation in $\mathbf{B}_1$ with a representation of solution given by the theorem. Inserting $\hat{\mathbf{B}}_1$ in (2.1) and (2.4) and doing some calculations we obtain

$$A'_r \Sigma^{-1} P_2 \left(X - \sum_{i=2}^{m} A_i \mathbf{B}_i C_i \right) C'_r = 0, \qquad r = 2, 3, ..., m \tag{2.6}$$

$$\Sigma = S_2 + P_2 \left(XC'_2(C_2C'_2)^- C_2 - \sum_{i=2}^{m} A_i \mathbf{B}_i C_i \right) \times \left(XC'_2(C_2C'_2)^- C_2 - \sum_{i=2}^{m} A_i \mathbf{B}_i C_i \right)' P'_2. \tag{2.7}$$

For $r = 2$, (2.6) implies that

$$A'_2 \Sigma^{-1} P_2 \left(XC'_2(C_2C'_2)^- C_2 - \sum_{i=2}^{m} A_i \mathbf{B}_i C_i \right) C'_2 = 0. \tag{2.8}$$

Using Lemma 2.1(ii) again leads to the equation

$$A'_2 S_2^{-1} P_2 \left(XC'_2(C_2C'_2)^- C_2 - A_2 \mathbf{B}_2 C_2 - \sum_{i=3}^{m} A_i \mathbf{B}_i C_i \right) HC'_2 = 0, \tag{2.9}$$

where H, although differing from H in (2.5), is p.d. Thus, utilizing Lemma 2.2, $P'_2 S_2^{-1} P_2 = S_2^{-1} P_2$ and (2.9) is of the same form as (2.5) showing the representation of solution in the theorem for $\hat{\mathbf{B}}_2$ to be valid. Proceeding in the same manner we obtain solutions for the other parameters. ∎

Note that in the above theorem we have just shown the representations of the estimators to be solutions to the likelihood equations and not that they are maximum likelihood estimators. However, it is easy to see that the representations given in the theorem maximize the likelihood. Just note that ($|\cdot|$ stands for the determinant)

$$\begin{aligned} &\left| S_{r-1} + P_r \left(X - \sum_{i=r}^{m} A_i \mathbf{B}_i C_i \right) C'_{r-1}(C_{r-1}C'_{r-1})^- C_{r-1} \left(X - \sum_{i=r}^{m} A_i \mathbf{B}_i C_i \right)' P'_r \right| \\ &\geqslant \left| S_r + P_{r+1} \left(X - \sum_{i=r+1}^{m} A_i \mathbf{B}_i C_i \right) C'_r(C_rC'_r)^- C_r \right. \\ &\quad \left. \times \left(X - \sum_{i=r+1}^{m} A_i \mathbf{B}_i C_i \right)' P'_{r+1} \right| \\ &\geqslant |n\hat{\Sigma}|, \qquad r = 2, 3, ..., m \end{aligned} \tag{2.10}$$

and equality holds if and only if $\hat{\mathrm{B}}_r$ and $\hat{\Sigma}$ satisfy Theorem 2.1. We also obtain that $\hat{\Sigma}$ always is uniquely estimated. Some details of showing (2.10) can be found in von Rosen [20], where a modification of an approach in Srivastava and Khatri [26] is presented.

Observe that solving the likelihood equations is divided into m steps. In each step, what might be expected from univariate linear models, a proper projection is made, which is built up by aid of P_r and $C_r'(C_rC_r')^- C_r$, and in each step an inner product defined by Σ is estimated with the help of S_r. Some comments on these aspects for the MLNM(ABC) are given by von Rosen [20].

The $\hat{\mathrm{B}}$'s are given in a recursive formula which is fairly useful. In order to give the expressions in a non-recursive manner one has to impose some kind of uniqueness condition such as $\sum_{i=r+1}^m A_i\hat{\mathrm{B}}_iC_i$ to be unique. Otherwise expressions given in a non-recursive way are rather hard to interpret. However, without any further assumptions $P_r\sum_{i=r}^m A_i\hat{\mathrm{B}}_iC_i$ is always unique and the next theorem gives, in a non-recursive manner, the expression for it.

THEOREM 2.2. *For the* $\hat{\mathrm{B}}$*'s given in Theorem* 2.1,

$$P_r\sum_{i=r}^m A_i\hat{\mathrm{B}}_iC_i=\sum_{i=r}^m (I-T_i)\,XC_i'(C_iC_i')^- C_i.$$

Proof. Since $P_r'S_r^{-1}P_r=P_r'S_r^{-1}$ (Lemma 2.2) implies that $(I-T_r)=(I-T_r)\,P_r$,

$$\begin{aligned}P_r\sum_{i=r}^m A_i\hat{\mathrm{B}}_iC_i&=(I-T_r)\,XC_r'(C_rC_r')^- C_r-(I-T_r)\sum_{i=r+1}^m A_i\hat{\mathrm{B}}_iC_i\\&\quad+P_r\sum_{i=r+1}^m A_i\hat{\mathrm{B}}_iC_i\\&=(I-T_r)\,XC_r'(C_rC_r')^- C_r+T_rP_r\sum_{i=r+1}^m A_i\hat{\mathrm{B}}_iC_i\\&=(I-T_r)\,XC_r'(C_rC_r')^- C_r+P_{r+1}\sum_{i=r+1}^m A_i\hat{\mathrm{B}}_iC_i. \quad\blacksquare\end{aligned}$$

A useful application of this theorem is when estimating the mean structure.

COROLLARY 2.1. $\widehat{E[X]}=\sum_{i=1}^m (I-T_i)\,XC_i'(C_iC_i')^- C_i$.

3. Linear Restrictions on B in the MLNM(ABC)

This section treats some useful applications of the MLNM($\sum_{i=1}^{m} A_i \mathrm{B}_i C_i$). First, we will see that the MLNM(ABC) when

$$D_1 \mathrm{B} E_1 = 0, \qquad D_2 \mathrm{B} E_2 = 0, \tag{3.1}$$

under some restrictions on the matrices, can be written as a MLNM($\sum_{i=1}^{m} A_i \mathrm{B}_i C_i$). The following lemma gives some algebraic results which this section will rest on.

Lemma 3.1. (i) *A representation of solution to* (3.1) *is given by*

$$\mathrm{B} = T_1 Z_1 E_2^{0\prime} + T_2 Z_2 (E_1 : E_2)^{0\prime} + T_3 Z_3 E_1^{0\prime} + T_4 Z_4,$$

where the Z's are arbitrary matrices and T_1–T_4 are any matrices satisfying

$$\mathbf{R}(T_1) = \mathbf{R}(D_1' : D_2') \cap \mathbf{R}(D_1')^{\perp}, \qquad \mathbf{R}(T_2) = \mathbf{R}(D_1') \cap \mathbf{R}(D_1')$$

$$\mathbf{R}(T_3) = \mathbf{R}(D_1' : D_2') \cap \mathbf{R}(D_2')^{\perp}, \qquad \mathbf{R}(T_4) = \mathbf{R}(D_1' : D_2')^{\perp}.$$

(ii) *A representation of solution to*

$$D_i \mathrm{B} E_i = 0, \qquad i = 1, 2, ..., s$$

when $\mathbf{R}(E_s) \subseteq \mathbf{R}(E_{s-1}) \subseteq \cdots \subseteq \mathbf{R}(E_1)$ *holds is given by* (*set* $H_i = D_1' : D_2' : \cdots : D_i'$)

$$\mathrm{B} = H_s^0 Z_1 E_s' + \sum_{i=1}^{s-1} H_i^0 Z_{i+1} (E_i^0 : E_{i+1})^{0\prime} + Z_{s+1} E_1^{0\prime}$$

or

$$\mathrm{B} = H_s^0 Z_1 + \sum_{i=2}^{s} (H_{i-1} : H_i^0)^0 Z_i E_i^{0\prime} + D_1' Z_{s+1} E_1^{0\prime},$$

where the Z's are arbitrary matrices.

Proof. The proof of (i) is given by von Rosen [22]. In order to verify (ii) we use, for notational convenience, instead of $\mathbf{R}(E_i)$, $\mathbf{R}(D_i')$, and $\mathbf{R}(H_i)$, $\mathbf{E}_i$, $\mathbf{D}_i'$, and $\mathbf{H}_i$, respectively. The tensor product is denoted $\otimes$ and, for example, the tensor product of $\mathbf{E}_i$ and $\mathbf{D}_i'$ is denoted $\mathbf{E}_i \otimes \mathbf{D}_i'$. By rewriting $D_i \mathrm{B} E_i$ into a vectorized forms follows that (ii) is true if it can be shown that

$$\begin{aligned}\left(\sum_{i=1}^{s} \mathbf{E}_i \otimes \mathbf{D}_i'\right)^{\perp} &= \mathbf{E}_s \otimes \mathbf{H}_s^{\perp} + \sum_{i=1}^{s-1} (\mathbf{E}_i \cap \mathbf{E}_{i+1}^{\perp} \otimes \mathbf{H}_i^{\perp}) + \mathbf{E}_1^{\perp} \otimes \mathbf{V}_2 \\ &= \mathbf{V}_1 \otimes \mathbf{H}_s^{\perp} + \sum_{i=2}^{s} (\mathbf{E}_i^{\perp} \otimes \mathbf{H}_i \cap \mathbf{H}_{i-1}^{\perp}) + \mathbf{E}_1^{\perp} \otimes \mathbf{D}_1', \end{aligned} \tag{3.2}$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ stand for the whole space for $\mathbf{E}_i$ and $\mathbf{D}_i'$, respectively (i.e., $\mathbf{V}_1 = \mathbf{E}_i + \mathbf{E}_i^{\perp}$, $\mathbf{V}_2 = \mathbf{D}_i' + (\mathbf{D}_i')^{\perp}$). By applying the orthomodular law (see Nordström and von Rosen [16]) we get

$$\mathbf{E}_i = \mathbf{E}_s + \sum_{j=i}^{s-1} \mathbf{E}_j \cap \mathbf{E}_{j+1}^{\perp}, \qquad i = 1, 2, ..., s-1,$$

and thus

$$\sum_{i=1}^{s} \mathbf{E}_i \otimes \mathbf{D}_i' = \mathbf{E}_s \otimes \mathbf{H}_s + \sum_{i=1}^{s-1} \sum_{j=i}^{s-1} \mathbf{E}_j \cap \mathbf{E}_{j+1}^{\perp} \otimes \mathbf{D}_i'$$
$$= \mathbf{E}_s \otimes \mathbf{H}_s + \sum_{i=1}^{s-1} \mathbf{E}_i \cap \mathbf{E}_{i+1}^{\perp} \otimes \mathbf{H}_i' \tag{3.3}$$

which obviously is orthogonal to the first statement in (3.2). Moreover, summing the first statement in (3.2) with (3.3) gives us the whole space. Hence, an orthogonal complement to $\sum_{i=1}^{s} \mathbf{E}_i \otimes \mathbf{D}_i'$ has been found. The second statement in (3.2) follows by straightforward manipulations and by noting that

$$\mathbf{H}_{i-1}^{\perp} = \mathbf{H}_i^{\perp} + \mathbf{H}_i \cap \mathbf{H}_{i-1}^{\perp}. \quad \blacksquare$$

Considering the MLNM(ABC) when (3.1) holds, we obtain from Lemma 3.1(i) that we, equivalently, can work with a model with mean structure defined by

$$E[X] = AT_1\theta_1 E_2^{0\prime}C + AT_2\theta_2(E_1 : E_2)^{0\prime}\, C + AT_3\theta_3 E_1^{0\prime}C + AT_4\theta_4 C, \tag{3.4}$$

where the θ's are new parameters. Note that the mean structure presented in (3.4) may not belong to the class of MLNM($\sum_{i=1}^{4} A_i \mathbf{B}_i C_i$) without any further assumptions since $\mathbf{R}(C'E_1^0) \subseteq \mathbf{R}(C'E_2^0)$ or $\mathbf{R}(C'E_2^0) \subseteq \mathbf{R}(C'E_1^0)$ do not have to hold. Therefore, in order to utilize the results of the previous section some conditions have to be imposed and it seems natural to suppose that at least one of the following four conditions is satisfied:

$$\begin{aligned} AT_1 &= 0, \qquad \mathbf{R}(C'E_2^0) \subseteq \mathbf{R}(C'E_1^0) \\ AT_3 &= 0, \qquad \mathbf{R}(C'E_1^0) \subseteq \mathbf{R}(C'E_2^0). \end{aligned} \tag{3.5}$$

By symmetry it follows that we only have to consider (3.5). Moreover, note that $\mathbf{R}(D_2') \subseteq \mathbf{R}(D_1')$ implies $AT_1 = 0$ and $\mathbf{R}(E_1) \subseteq \mathbf{R}(E_2)$ implies $\mathbf{R}(C'E_2^0) \subseteq \mathbf{R}(C'E_1^0)$.

THEOREM 3.1. *Suppose that for the* MLNM(ABC) (3.1) *holds. If* $\mathbf{R}(C'E_2^0) \subseteq \mathbf{R}(C'E_1^0)$ *or* $AT_1 = 0$ *the maximum likelihood estimators are obtainable from a* MLNM($\sum_{i=1}^{m} A_i \mathbf{B}_i C_i$).

Note that Σ is always uniquely estimated and if $\rho(A)=q$ as well as $\rho(C)=k$ in the MLNM(ABC) B is also always uniquely estimated.

Although estimators for the parameters in the MNLM(ABC) when constraints like those in (3.1) exist have not hitherto been presented, likelihood ratio tests have been derived under conditions slightly stronger than those presented in this section (see Banken [3, 4] and Kariya [11]). However, since $\hat{\Sigma}$ is unique, whether or not there exist constraints, we can immediately set up the likelihood ratio criterion. In comparison to Banken [3, 4] and Kariya [11] we can express the criterion in the original matrices. Note that we do not have to assume any testability conditions, since $\hat{\Sigma}$ is unique. Of course, the tests are only meaningful if the constraints have some impact on the estimator for Σ but in order to write down the likelihood ratio criterion one does not have to consider this.

COROLLARY 3.1. *Suppose that for the* MLNM(ABC), $D\mathrm{B}E=0$.

(i) $\mathrm{B}=(D')^0\theta_1+D'\theta_2E^{0'}$

(ii) $E[X]=A(D')^0\theta_1C+AD'\theta_2E^{0'}C$

(iii) *Maximum likelihood estimators are obtainable from a* MLNM($\sum_{i=1}^m A_i\mathrm{B}_iC_i$).

Proof. The statements follow, for instance, from Lemma 3.1(i), (3.4), and Theorem 3.1 if we assume $D_2=0$. ∎

Tubbs *et al.* [27] considered the problem of estimating B under the restriction $D\mathrm{B}E=0$ in the MLNM(ABC). Sometimes it has been alleged that the estimator proposed by Tubbs *et al.* is a maximum likelihood estimator, but the estimator is not in agreement with ours obtained by combining Corollary 3.1 and Theorem 2.1. Hence, their estimator is not a maximum likelihood estimator. Moreover, a likelihood ratio test for $D\mathrm{B}E=0$ in the MLNM(ABC) can be constructed by aid of Corollary 3.1 and it follows, after quite a lot of matrix manipulations (von Rosen [20]), that this test is, of course, identical to Khatri's [12] (see also Baksalary and Kala [2]). In the next, Theorem 3.1 is extended.

THEOREM 3.2. *Suppose that for the* MLNM(ABC)

$$D_i\mathrm{B}E_i=0, \qquad i=1, 2, ..., s,$$

where

(i) $\mathbf{R}(E_s)\subseteq\mathbf{R}(E_{s-1})\subseteq\cdots\subseteq\mathbf{R}(E_1)$ *holds. Then (set* $H_i=D_1':D_2':\cdots:D_i'$)

$$E[X]=AH_s^0\theta_1C+A\sum_{i=2}^{s}(H_{i-1}:H_i^0)^0\,\theta_iE_i^{0'}C+AD_1'\theta_{s+1}E_1^{0'}C.$$

(ii) $\mathbf{R}(D_s') \subseteq \mathbf{R}(D_{s-1}') \subseteq \cdots \subseteq \mathbf{R}(D_1')$ *holds. Then* (*set* $G_i = E_1 : E_2 : \cdots : E_i$)

$$E[X] = AD_s'\theta_1 G_s^{0\prime}C + \sum_{i=1}^{s-1} A((D_i')^0 : D_{i+1}')^0 \theta_{i+1} G_i^{0\prime}C + A(D_1')^0 \theta_{s+1} C.$$

Proof. Both statements are obtained by aid of Lemma 3.1(ii). ∎

Note that the criterion $\mathbf{R}(C_m') \subseteq \mathbf{R}(C_{m-1}') \subseteq \cdots \subseteq \mathbf{R}(C_1')$ in the MLNM($\sum_{i=1}^m A_i \mathrm{B}_i C_i$) is satisfied. Possible extensions of Theorem 3.1 and Theorem 3.2 to cover restrictions about B_i in the MLNM($\sum_{i=1}^m A_i \mathrm{B}_i C_i$) are also fairly easy to derive.

4. Missing Observations

This section will be devoted to the problem when missing observations exist. The MLNM($\sum_{i=1}^m A_i \mathrm{B}_i C_i$) will be the model under consideration and thus in line with the previous section we can obtain estimators when linear restrictions on B in the MLNM($A\mathrm{B}C$) exist. The work of Kleinbaum [13] seems to be one of the first in estimating and testing linear functions of B in the MLNM($A\mathrm{B}C$) under the assumption that some values are missing. However, if data consist of many missing values Kleinbaum's approach is difficult to apply. This was pointed out by Liski [14, 15]. As an alternative to Kleinbaum's approach Liski [15] presented work on growth curve data with missing values where the EM algorithm (for details of the algorithm see Dempster *et al.* [6]) was utilized when estimating the parameters. The EM algorithm, besides slow convergence when there is not a clear maximum, has several desirable qualities. For instance, the algorithm produces maximum likelihood estimators and is easy to program for linear normal models. Jennrich and Schluchter [10], when maximizing the likelihood, discussed several numerical aspects and adjoining algorithms to the EM algorithm. Srivastava [25] studied the likelihood equations. Both Jennrich and Schluchter's and Srivastava's approaches can be applied to the MLNM($\sum_{i=1}^m A_i \mathrm{B}_i C_i$) with missing data but, in order to use their methods effectively, Theorem 2.1 as well as our approach of solving the likelihood equations may be directing.

In this section we illustrate the missing data problem by applying the EM algorithm to the MLNM($\sum_{i=1}^m A_i \mathrm{B}_i C_i$) and in what follows some notations are needed. Let $\hat{\Theta}^k$ stand for the value of the parameters in the kth iteration, $E[\cdot \mid \cdot]$ signifies conditional expectation, $l(\cdot;\cdot)$ means the log likelihood, and X_0 denotes the elements in X which correspond to the nonmissing observations. The kth iteration of the EM algorithm is given by the following two steps:

(i) Calculate $E[l(X;\Theta)\mid X_0, \hat{\Theta}^{k-1}]$.

(ii) Let $\hat{\Theta}^k$ be the value of Θ which maximizes the expectation in (i).

An initial estimate of Θ ($\hat{\Theta}^0$) may be obtained by aid of the complete data set X_0. Moreover, let $\mu = E[X]$, $\hat{\Theta}^k = (\hat{\mu}^k, \hat{\Sigma}^k)$, and

$$\hat{X}^k = E[X\mid X_0, \hat{\Theta}^k], \qquad U^k = E[(X-\hat{X}^k)(X-\hat{X}^k)'\mid X_0, \hat{\Theta}^k].$$

Utilizing these notations we obtain

$$E[l(X;\mu,\Sigma)\mid X_0,\hat{\Theta}^k] = -\tfrac{1}{2}np\ln(2\pi) - \tfrac{1}{2}n\ln(|\Sigma|) - \tfrac{1}{2}\mathrm{tr}(\Sigma^{-1}((\hat{X}^k-\mu)(\hat{X}^k-\mu)' + U^k)). \quad (4.1)$$

Note that U^k is positive semidefinite and since the columns of X are normally distributed $\hat{X}^k$ and U^k are easily computed. In order to obtain the $(k+1)$th iteration, (4.1) is maximized with respect to μ and Σ when $\mu = \sum_{i=1}^m A_i \mathrm{B}_i C_i$ and $\mathbf{R}(C_m') \subseteq \mathbf{R}(C_{m-1}') \subseteq \cdots \subseteq \mathbf{R}(C_1')$ hold. However, since the likelihood equations in the $(k+1)$th iteration equal

$$A_r'\Sigma^{-1}\left(\hat{X}^k - \sum_{i=1}^m A_i\mathrm{B}_iC_i\right)C_r' = 0, \qquad r = 1, 2, ..., m,$$

$$n\Sigma = U^k + \left(\hat{X}^k - \sum_{i=1}^m A_i\mathrm{B}_iC_i\right)\left(\hat{X}^k - \sum_{i=1}^m A_i\mathrm{B}_iC_i\right)'$$

it follows from the proof of Theorem 2.1 that the next theorem is established.

THEOREM 4.1. *The $(k+1)$th EM iteration for the* MLNM($\sum_{i=1}^m A_i\mathrm{B}_iC_i$) *with missing observations is given by replacing* $\hat{\mathrm{B}}_r$, $r = 1, 2, ..., m$, $\hat{\Sigma}$, X, *and* K_1 *in Theorem* 2.1 *by* $\hat{\mathrm{B}}_r^{k+1}$, $\hat{\Sigma}^{k+1}$, $\hat{X}^k$, *and* $\hat{X}^k(I - C_1'(C_1C_1')^- C_1)(\hat{X}^k)' + U^k$, *respectively.*

5. ML-Estimators for the MLNM($\sum_{i=1}^m A_i\mathrm{B}_iC_i + \mathrm{B}_{m+1}C_{m+1}$)

Here we will extend the works by Elswick [7], Chinchilli and Elswick [5] and von Rosen [20, 23]. The MLNM($\sum_{i=1}^m A_i\mathrm{B}_iC_i + \mathrm{B}_{m+1}C_{m+1}$) is especially applicable to growth curve data when covariate variables exist. For completeness it is noted that in the definition of the model X is as in the previous sections but the mean structure equals

$$E[X] = \sum_{i=1}^m A_i\mathrm{B}_iC_i + \mathrm{B}_{m+1}C_{m+1},$$

where the B's are the parameters, the A's and C's known design matrices and $\mathbf{R}(C'_m) \subseteq \mathbf{R}(C'_{m-1}) \subseteq \cdots \subseteq \mathbf{R}(C'_1)$. We stress that there is no restriction on $\mathbf{R}(C'_{m+1})$.

THEOREM 5.1. *Let*

$$\bar{C}_i = C_i(I - C'_{m+1}(C_{m+1}C'_{m+1})^- C_{m+1}), \qquad r = 1, 2, ..., m,$$

$$P_r = T_{r-1}T_{r-2}T_{r-3} \times \cdots \times T_0, \qquad T_0 = I, \qquad r = 1, 2, ..., m+1,$$

$$T_i = I - P_iA_i(A'_iP'_iS_i^{-1}P_iA_i)^- A'_iP'_iS_i^{-1}, \qquad i = 1, 2, ..., m,$$

$$S_i = \sum_{j=1}^{i} K_j$$

$$K_j = P_jX\bar{C}'_{j-1}(\bar{C}_{j-1}\bar{C}'_{j-1})^- \bar{C}_{j-1}(I - \bar{C}'_j(\bar{C}_j\bar{C}'_j)^- \bar{C}_j)\,\bar{C}'_{j-1}(\bar{C}_{j-1}\bar{C}'_{j-1})^- \times \bar{C}_{j-1}X'P'_j, \qquad j = 2, 3, ..., m,$$

$$K_1 = X(I - (C'_{m+1} : C'_1)((C'_{m+1} : C'_1)'(C'_{m+1} : C'_1))^- (C'_{m+1} : C'_1)')\,X'.$$

Assuming S_1 to be p.d., representations of the maximum likelihood estimators for the $\mathrm{MLNM}(\sum_{i=1}^m A_i\mathbf{B}_iC_i + \mathbf{B}_{m+1}C_{m+1})$ *are given by*

$$\hat{\mathbf{B}}_r = (A'_rP'_rS_r^{-1}P_rA_r)^- A'_rP'_rS_r^{-1}\left(X - \sum_{i=r+1}^{m} A_i\hat{\mathbf{B}}_i\bar{C}_i\right)\bar{C}'_r(\bar{C}_r\bar{C}'_r)^- + (A'_rP'_r)^0 Z_{r1} + A'_rP'_rZ_{r2}\bar{C}_r^{0\prime}, \qquad r = 1, 2, ..., m,$$

$$\hat{\mathbf{B}}_{m+1} = \left(X - \sum_{i=1}^{m} A_i\hat{\mathbf{B}}_iC_i\right)C'_{m+1}(C_{m+1}C'_{m+1})^- + Z_{m+1}C_{m+1}^{0\prime}$$

$$n\hat{\Sigma} = \left(X - \sum_{i=1}^{m} A_i\hat{\mathbf{B}}_iC_i - \hat{\mathbf{B}}_{m+1}C_{m+1}\right)\left(X - \sum_{i=1}^{m} A_i\hat{\mathbf{B}}_iC_i - \hat{\mathbf{B}}_{m+1}C_{m+1}\right)'$$

$$= S_m + P_{m+1}X\bar{C}'_m(\bar{C}_m\bar{C}'_m)^- \bar{C}_mX'P'_{m+1},$$

where the Z's are arbitrary matrices.

Proof. We are going to show that the theorem follows from the proof of Theorem 2.1. The likelihood equations for the $\mathrm{MLNM}(\sum_{i=1}^m A_i\mathbf{B}_iC_i + \mathbf{B}_{m+1}C_{m+1})$ equal

$$A'_r\Sigma^{-1}\left(X - \sum_{i=1}^{m} A_i\mathbf{B}_iC_i - \mathbf{B}_{m+1}C_{m+1}\right)C'_r = 0, \qquad r = 1, 2, ..., m, \tag{5.1}$$

$$\left(X - \sum_{i=1}^{m} A_i\mathbf{B}_iC_i - \mathbf{B}_{m+1}C_{m+1}\right)C'_{m+1} = 0 \tag{5.2}$$

$$n\Sigma = \left(X - \sum_{i=1}^{m} A_i\mathbf{B}_iC_i - \mathbf{B}_{m+1}C_{m+1}\right)\left(X - \sum_{i=1}^{m} A_i\mathbf{B}_iC_i - \mathbf{B}_{m+1}C_{m+1}\right)'. \tag{5.3}$$

Note that (5.2) is a linear equation in B_{m+1} and hence $\hat{\mathrm{B}}_{m+1}$ is obtained. Inserting $\hat{\mathrm{B}}_{m+1}$ in (5.1) and (5.3) gives the equations

$$A_r' \Sigma^{-1} \left(X - \sum_{i=1}^{m} A_i \mathrm{B}_i \bar{C}_i \right) \bar{C}_r' = 0, \qquad r = 1, 2, ..., m,$$

$$n\Sigma = S_1 + \left(X\bar{C}_1'(\bar{C}_1 \bar{C}_1')^- \bar{C}_1 - \sum_{i=1}^{m} A_i \mathrm{B}_i \bar{C}_i \right)$$
$$\times \left(X\bar{C}_1'(\bar{C}_1 \bar{C}_1')^- \bar{C}_1 - \sum_{i=1}^{m} A_i \mathrm{B}_i \bar{C}_i \right)'$$

which are identical to (2.1) and (2.4). ∎

Remark. Note that the theorem is still valid if instead of $\mathbf{R}(C_m') \subseteq \mathbf{R}(C_{m-1}') \subseteq \cdots \subseteq \mathbf{R}(C_1')$ the weaker condition $\mathbf{R}(\bar{C}_m') \subseteq \mathbf{R}(\bar{C}_{m-1}') \subseteq \cdots \subseteq \mathbf{R}(\bar{C}_1')$ is used.

References

1. Anderson, T. W., and Olkin, I. (1985). Maximum-likelihood estimation of the parameters of a multivariate normal distribution. *Linear Algebra Appl.* **70** 147–171.
2. Baksalary, J. K. and Kala, R. (1977). Sums of squares and products matrices for a non-full ranks hypothesis in the model of Potthoff and Roy. *Math. Operationsforsch. Statist.* **8** 459–465.
3. Banken, L. (1984a). *On the Reduction of the General MANOVA Model.* Tech. Report, University of Trier, Trier, West Germany.
4. Banken, L. (1984b). *Eine Verallgemeinerung des GMANOVA-Modells.* Dissertation, University of Trier, Trier, West Germany.
5. Chinchilli, V. M., and Elswick, R. K. (1985). A Mixture of the MANOVA and GMANOVA models. *Commun. Statist.-Theor. Methods* **14** 3075–3089.
6. Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B* **39** 1–38.
7. Elswick, R. K. (1985). *The Missing Data Problem as Applied to the Extended Version of the GMANOVA Model.* Dissertation, Virginia Commonwealth University, Richmond, VA.
8. Gleser, L. J., and Olkin, I. (1970). Linear models in multivariate analysis. In *Essays in Probability and Statistics* (R. C. Bose, I. M. Chakravarti, P. C. Mahalanobis, C. R. Rao, and K. J. C. Smith, Eds.), pp. 267–292. University of North Carolina Press, Chapel Hill, NC.
9. Grizzle, J. E., and Allen, D. M. (1969). Analysis of growth and dose response curves. *Biometrics* **25** 357–381.
10. Jennrich, R. I., and Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics* **42** 805–820.
11. Kariya, T. (1985). *Testing in the Multivariate General Linear Model.* Kinokuniya, New York.
12. Khatri, C. G. (1966). A note on a MANOVA model applied to problems in growth curve. *Ann. Inst. Statist. Math.* **18** 75–86.

13. KLEINBAUM, D. G. (1973). A generalization of the growth curve model which allows missing data. *J. Multivariate Anal.* **3** 117–124.
14. LISKI, E. P. (1984). *A Bayesian Approach to Missing Data Estimation in Growth Curves.* Technical Report A 135. Dept. of Math. Sciences, University of Tampere, Tampere, Finland.
15. LISKI, E. P. (1985). Estimation from incomplete data in growth curves models. *Commun. Statist.-Simul. Comput.* **14** 13–27.
16. NORDSTRÖM, K., AND VON ROSEN, D. (1987). Algebra of subspaces with applications to problems in statistics. In *Proceedings, Second International Tampere Conference in Statistics* (T. Pukkila and S. Puntanen, Eds.), pp. 603–614. University of Tampere, Tampere, Finland.
17. POTTHOFF, R. F., AND ROY, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* **51** 313–326.
18. RAO, C. R. (1965). The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika* **52** 447–458.
19. RAO, C. R. (1966). Covariance adjustment and related problems in multivariate analysis. In *Multivariate Analysis* (P. R. Krishnaiah, Ed.), pp. 87–103. Academic Press, New York.
20. VON ROSEN, D. (1984). *Maximum Likelihood Estimates in Multivariate Linear Normal Models with Special References to the Growth Curve Model.* Research Report 135, Dept. of Math. Statist., University of Stockholm, Stockholm, Sweden.
21. VON ROSEN, D. (1985). *Multivariate Linear Normal Models with Special References to the Growth Curve Model.* Dissertation, University of Stockholm, Stockholm, Sweden.
22. VON ROSEN, D. (1986a). Some results on homogenous matrix equations. Submitted for publication.
23. VON ROSEN, D. (1986b). Combining MANOVA and GMANOVA, In *Proceedings, IVth International Conference on Multivariate Statistical Inference. Institute of Econometrics and Statistics, University of Łódź, Łódź, Poland.*
24. SEBER, G. A. F. (1985). *Multivariate Observations.* Wiley, New York.
25. SRIVASTAVA, M. S. (1985). Multivariate data with missing observations. *Commun. Statist.-Theor. Methods* **14** 775–792.
26. SRIVASTAVA, M. S., AND KHATRI, C. G. (1979). *An Introduction to Multivariate Statistics.* North-Holland, New York.
27. TUBBS, J. D., LEWIS, T. O., AND DURAN, B. S. (1975). A note on the analysis of the MANOVA model and its application to growth curves. *Commun. Statist.* **4** 643–653.
28. WOOLSON, R. F., AND LEEPER, J. D. (1980). Growth curve analysis of complete and incomplete longitudinal data. *Commun. Statist.-Theor. Methods* **9** 1491–1513.